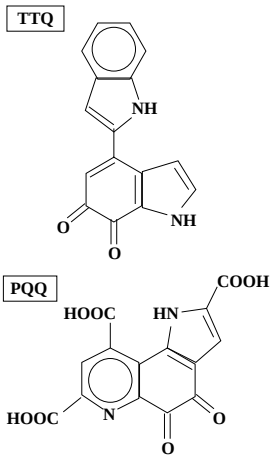
1 Innovative discoveries and experiences

Biochemical pathways involving electron transfer offer an appealing route for the development of bioanalytic sensors. Among the simplest such pathways to characterize experimentally is a series of electron transfer reactions that provide the mechanism for dehydrogenation of methanol and methylamine. The enzymes that promote these processes have well-established structures, including data on the influence of the physiological environment [1]. This is in contrast to the membrane-integrated proteins that direct many other electron transfer processes. Furthermore, the methanol and methylamine dehydrogenation reactions are based on electron donation from quinone cofactors, which are unusual in offering a tractably small molecular subunit for computational modeling of the reaction. The cofactors that form the foundation of these reactions are drawn in Fig. 1.

The proposed study is a computational investigation of the chemical reaction mechanisms for the enzyme activity of methanol dehydrogenase (MEDH) and methylamine dehydrogenase (MADH), focusing on the activity of their respective cofactors: pyrroloquinoline quinone (PQQ) and tryptophan tryptophylquinone (TTQ). The computational methods to be used have been shown by the PI's research group to be effective in the prediction of reaction properties of small linear and cyclic organic molecules, similar to the active sites in the MEDH and MADH cofactors. In particular, that previous work has focused on the structures and dynamics of free radicals with conjugated π-electron systems, which account for the most challenging chemical species to be faced in the proposed work.

Figure 1: The chemical structures of cofactors PQQ and TTQ



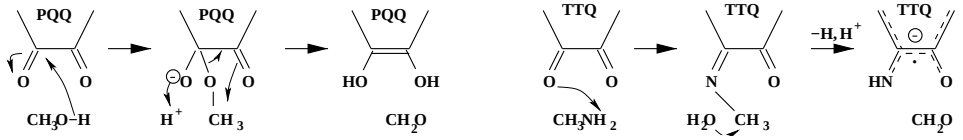
The chemical reactions that will be described are

$$\mathrm{RCH_2OH} + 2\mathrm{A_{ox}} \xrightarrow{\mathrm{MEDH}} \mathrm{RCHO} + 2\mathrm{A^-_{red}} + 2\mathrm{H^+} \quad (1)$$

$$\mathrm{RCH_2NH_2} \xrightarrow{\mathrm{MADH}} \mathrm{RCHO} + \mathrm{NH_3}, \quad (2)$$

where $\mathrm{A_{ox}}$ and $\mathrm{A_{red}}$ are oxidized and reduced forms of an electron acceptor. These reactions have been targeted by several diverse analytic and kinetic studies, most recently an investigation of effects on the electron transfer rate by engineering changes to the protein electron acceptor [2]. The mechanisms illustrated schematically in Fig. 2 are fairly well established by experiment. However, the reactions have not yet been probed by rigorous computational methods. Given the size of these molecules and the nature of the interactions involved, the proposed computational investigation is likely to be qualitatively accurate, but the precision possible from the in silico studies remains to be determined. If rate constants for the kinetics of the elementary steps can be predicted even to within an order of magnitude, this will help to establish quantum computational chemistry as a powerful tool in the elucidation of biochemical pathways.

Figure 2: Abbreviated scheme for PQQ and TTQ dehydrogenations, showing only the activity at the quinone site. Arrows indicate redistribution of electron pairs. Regeneration of the quinone catalyst occurs by reaction with water (and subsequent loss of NH_3 for TTQ). After Figs. 7 and 15 in Ref. 1.



The principal innovation in the proposed research is the application of high-level ab initio quantum chemical methods – which predict chemical properties using little other than the fundamental laws of quantum physics – to an important biochemical process. Computational modeling of biochemistry is often limited to very approximate, classical methods by the enormous numbers of atoms and the importance of large-scale structure. Such classical studies are invaluable to an understanding of protein folding and molecular docking, but are also fatally crude when applied to problems that require an accurate depiction of the electron distribution, and especially those that rely on identification of the transition state structures that act as the gate between reactants and products.

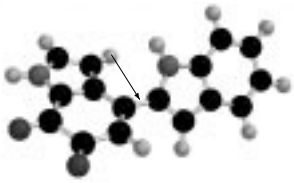
The PQQ and TTQ reactions involve sufficiently small reactive components that ab initio methods are challenging but feasible. Despite the wealth of experimental data, only a single, low-level computational study of one of these systems exists in the literature [3]. In the proposed study, the electronic structure will be modeled using a linear combination of roughly 350 Gaussian functions, which are optimized to solve the Schrödinger equation using coupled cluster (CC) methods [4]. The PI's current study of ring-opening reactions in cyclocarbinyl and previous work on cyclopentenyl [5] radicals have found this combination of methods to allow the determination of experimental geometries to within the limits of available experimental error, and energies of transition state structures are found to predict the activation barrier and pre-exponential factor of the cyclocarbinyl ring-opening rate constant to within 10%.

However, this level of accuracy in the kinetic properties is achieved only at the very demanding CCSD level of theory, following corrections for the vibrational motion of the atoms. The PQQ and TTQ calculations will require four times as many Gaussian functions as the carbinyl radical to describe the electron distribution, and the computational resources required by an energy calculation at a single geometry vary as this number to roughly the 1.5 power. Furthermore, because both PQQ and TTQ have about three times as many atoms as carbinyl and cyclopentenyl, the geometry optimizations and frequency calculations will test proportionately more geometries. Overall, each calculation on PQQ and TTQ species is expected to require about 25 times the cpu clock cycles and eight times the disk space and minimum memory required in the PI's previous studies. A geometry optimization at this level of theory is likely to be a two-week, single-processor calculation on the requested equipment.

The PQQ reaction outlined in Fig. 2 involves at least the three species shown plus two

transition states, and it will be necessary to consider non-concerted processes that discretize the several electron-pushing arrows in the second structure. Altogether, at least eight geometries will be optimized for the reaction, followed by consideration of the effects from chemical modifications that have been experimentally characterized. Vibrational frequency analysis will be carried out at each of these geometries to assess the vibrational contributions to the enthalpies and entropies of reaction, and energies will be recalculated at the optimized geometries using the more accurate CCSD(T) method and larger basis sets to test for convergence of the results. Augmented basis sets may be required for an adequate treatment of the anion charge distributions, and these are likely to double or triple the time and disk space required for energy calculation. The PI has experience with the convergence of similar calculations from previous studies. Fig. 2 shows that the TTQ reaction system proceeds through a resonance-stabilized radical anion intermediate. The radical structure introduces new challenges to the project, because the delocalized unpaired electron can substantially complicate the analysis of the molecule's electronic structure. The PI has extensive experience with such systems, and has shown the level of theory and basis functions described above to be capable of accurate kinetic as well as thermodynamic predictions.

Figure 3: Low-level computed structure of TTQ showing the conformational twist about the single bond (see arrow) connecting the two bicyclic substructures.



The impact of hydrogen bonding to the cofactors from the enzyme residues that comprise the cofactor environment will also be assessed. In the case of the PQQ reaction, an additional influence is the proximity of a Ca^{2+} or other dication. The effects on observed spectra and enzyme activity of replacing the naturally occurring Ca^{2+} with strontium and barium ions have been experimentally measured [1], and we plan to model these effects computationally. Our ability to predict the effect on the spectra is an especially discriminating test of the quality of the computed electron distribution. Finally, both TTQ and PQQ have distinct conformations – geometries that differ only by facile rotations about one or more single bonds – which can significantly effect the energetics of the chemistry (Fig. 3), and a low-level computational search of the available conformations will also need to be carried out at each step.

2 Impact within the community

The immediate goal of the proposed research is to successfully model the reactions described above, with the expectation that workers in the local bioanalytics industry will then use our model to predict how these chemical properties can be modified for use in biochemical sensors. A larger goal is to establish the extent to which the software and hardware of computational quantum chemistry are now capable of *quantitatively* describing fundamental biochemical processes, particularly those involving small, discrete chemical subunits such as TTQ and PQQ. Quantum computational models for testing the effects of modifications

on biochemical pathways will constitute a powerful addition to the toolbox of the local biomolecular industry.

The funds requested would provide a fellowship for an undergraduate student to work on this project, and for the purchase of eight computer nodes to enhance the central linux computer cluster of the SDSU Computational Science Research Center (CSRC). Raymond Wight, a San Diego high school graduate and now a B.S. Physics candidate at SDSU with three years of research experience in our lab, will be the primary researcher on this project.

The CSRC provides computer and staff resources to about two dozen SDSU faculty, and is actively pursuing partnerships with local industry. The PI for this proposal was recently awarded an NSF instrumentation grant to establish the Center's primary linux computer cluster. The computer nodes requested in this proposal will be added to that cluster, and will revert to general use by all CSRC researchers (including the PI) at the conclusion of the proposed project. The CSRC cluster, scheduled to go on-line in mid-April 2003, was funded by NSF grant CHE-0216563 but at a reduced budget following the summer 2002 fiscal crisis. The system is therefore smaller than originally requested, and would be overburdened by the proposed project, given the need for multiple simultaneous computations. At least ten researchers have projects intended to run on the cluster when it becomes established, in addition to the PI's own projects in basic molecular physics. The proposed addition of eight nodes would provide sufficient resources on this system for the major thrust of the PQQ/TTQ studies, and would subsequently become an important asset to the CSRC infrastructure.

3 Performance targets

The following target schedule is planned, based on a start date of July 15, 2003:

- Sep. 2003: optimization of all PQQ and TTQ geometries, in the absence of weak bonding; testing of relative conformational energies.
- Dec. 2003: completion of vibrational analysis on all minimum energy structures; determination of effects of weak bonding to enzyme residues on charge distributions and energies.
- Feb. 2003: optimization of all transition state structures.
- Apr. 2003: completion of follow-through calculations (should it be prudent to pursue investigation of side reactions, low-lying electronic states, or other potential complicating factors); calculation of mass-dependencies and tunneling effects for comparison to experiment.
- May 2003: submission of manuscripts on the kinetic models for both systems to a peer-reviewed journal (e.g., *Biochemistry* or *J. Biol. Chem.*).

4 Budget narrative

Funds are requested for the following: (*i*) a fellowship for one undergraduate researcher (full-time through August 2003; half-time academic year); (*ii*) one month summer salary for the PI during initial project development; (*iii*) eight dual-processor computer nodes from Western Scientific to expand the funded 16-node Intel/Linux CSRC computer cluster at SDSU; (*iv*) assorted office and computer supplies, including one dual-processor Intel/Linux computer (without monitor) for dedicated use in the research lab, for visualization and test calculations. The proposed project and extensions of it are suitable for funding by NIH , which has funded most of most of the precious work on this topic. Early in the funding period of the present proposal, the PI will submit a proposal to NIH for a grant to support research on this and other biochemical systems that rely on small enzyme cofactors. Because the proposed research represents the first extension of the PI's research into biochemical applications, chances to secure such funding will certainly be enhanced by preliminary results obtained under the requested grant.

Capital funds are requested to provide eight computer nodes to support the research. These will be dedicated to the project during the funding period, but will be housed as part of the CSRC cluster. This will allow the PI to take advantage of the system administration support and hardware support of the computer cluster, and will allow any nodes unused during the project period to become available to other CSRC-affiliated researchers. At the conclusion of the chief computational phase of the project (estimated April 2003), the eight nodes will become generally available to all CSRC users.

A dedicated workstation is also requested for the research lab, for the researcher on this project to visualize the structures and electron distributions resulting from the proposed calculations, and to provide a local computer for carrying out the many shorter low-level calculations (optimizing initial geometries, searching comformational geometries, predicting spectra based on electron distributions for the chemical intermediates, etc.) that contribute to the overall project.

References

[1] Davidson, V. L. *Adv. Prot. Chem.* **2000,** *58*, 95–140.

[2] Sun, D. P.; Davidson, V. L. *Biochemistry* **2003,** *42*, 1772–1776.

[3] Zheng, Y.-J.; Bruice, T. C. *Proc. Natl. Acad. Sci.* **1997,** *94*, 11881–11886.

[4] Purvis, III, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982,** *76*, 1910–1918.

[5] Martinez, C.; Cooksy, A. L. *J. Org. Chem.* **2002,** *67*, 2295–2302.